UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/091,689	03/06/2002	Bojana Gajic	2001-0105	7430

26652 7590 12/29/2005

AT&T CORP.
P.O. BOX 4110
MIDDLETOWN, NJ 07748

EXAMINER	
PIERRE, MYRIAM	
ART UNIT	**PAPER NUMBER**
2654	

DATE MAILED: 12/29/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/091,689	GAJIC ET AL.	
	Examiner	Art Unit	
	Myriam Pierre	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *21 November 2005*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-26* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-26* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. Applicants arguments, filed 11/21/2005 regarding the Office Action of 09/22/2005. Applicant amends claims 1, 2, 4-5, 9, 12-14, and 21-26, added claims 21-26.

Response to Arguments

2. Applicant's arguments with respect to claims 1, 4, 5 and 9-19 and 2, 6, 8 and 20 have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1, 5, 9-19, 22, 24, and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gong (6,418,411) in view of Digalakis et al. (5,864,810), in further view of DeVries (6,289,309).

As to claim 1 Gong teaches

a method of dynamically re-configurable speech recognition comprising:

determining an identity of a speaker based, at least in part, on a user identifier (col. 3 lines 13-18)

repeatedly (continually) determining parameters of a background model based on sampled information collected at periodic time interval (Fig. 2, 0.3 delay, col. 2 lines 35-45) during a received voice request {incoming utterance} (produce an adapted model based on inputs from on-line noise estimations (background adaptation) and one-time adaptation (transducer model), incoming utterance, col. 1, lines 42, 59-63, col. 2, lines 44-50 and Fig. 1, elements 11 & 20).

determining parameters (noise sample and utterance) of a transducer model (microphone or speaker) (Fig. 2; col. 5 lines 24-25);

adapting a speech recognition model based on user-specific transformations corresponding to the determined identity of the speaker (col. 3 lines 5-20) and on at least one of the background model (background noise) (Fig. 1 element 21 recognition, element 19 background noise, and col. 2 lines 59-61 steps 4-5);

Gong does not teach rescoring ASR.

However, Digalakis et al. do teach

re-scoring automatic speech recognition using the speech recognition model comprising:

generating word lattices representative of speech utterances in he received voice request (col. 11, lines 40-44);

concatenating the word lattices into a single concatenated lattice (sentence hypothesis necessarily implies word lattices, co. 13, lines 45-46);

applying at least one language model (language model) to the single concatenated lattice in order to determine word lattice inter-relationships (col. 13, lines 38-46); and

determining information in the received voice request based on he re-score results of the speech recognition model (rescoring the N-best sentence hypothesis, col. 13, lines 45-46); and

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify Gong 's method of speaker adaptation by re-scoring ASR that generates and links words in order to improve recognition performance for non-native speakers of American English, as taught by Digalakis et al., col. 13, lines 29-30.

Gong in view of Digalakis et al. does not explicitly teach adjusting the periodic time interval based on the determined changes in the sample.

However, DeVries et al. do teach

adjusting the periodic time interval based on the determined changes in the sampled information (col. 6 lines 10-24).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong in view of Digalakis et al. speaker identification because DeVries et al. teach that would produce noise tracking system that determines the effective time window in real time, so as to adapt to environmental changes in noise. (DeVries, col. 6 lines 18-23).

As to claim 5 Gong teaches

A system of dynamically re-configurable speech recognition comprising:

a background model estimation circuit for repeatedly determining a background model at a periodic time interval during a voice request based, at least in part, on estimated background parameters based on collected sampled information (background noise is recorded and estimated, col. 2, lines 43-44; and col. 5 lines 24-34; the background noise model is implemented via a

microphone and/or transducer which necessarily has the circuit for repeat determination of background noise as is needed in a noisy car environment);

a transducer model estimation circuit for determining a transducer model of the voice request based, at least in part, on estimated transducer parameters (col. 2, lines 35-44; and col. 5 lines 20-34);

a background model adaptation circuit and a transducer model adaptation circuit for determining an adapted speech recognition model based on a speech recognition model and at least on of the background model (col. 5 lines 5-10)

a lattice concatenation circuit that concatenates at least two speech lattices based on speech utterances in the received voice request into a signal lattice (col. 5 lines 5-34; speech recognition necessarily has a lattice link in order to determine the differences between speech and noise)

Gong does not explicitly teach adapting the controller based on user identification.

However, Digalakis et al. do teach

a controller that applies at least one language model to the signal concatenated lattice to determine relationships between lattices (col. 11 lines 20-26 and col. 6 lines 10-24).

the controller is adapted to determine an identify of a speaker based, at least in part on a user identified and to apply user-specific transformations, corresponding to the identity of the speaker, to the speech recognition model (Fig. 1-2 and col. 3 lines 20-25 and 43-47).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the method of Gong's speaker identification into the system of Digalakis et al. because Digalakis et al. teach that would improve performance and robustness of a speech

recognition system that is adapted to the speaker, and to the channel and the task (Digalakis, col. 2 lines 24-26).

Gong in view of Digalakis et al. does not explicitly teach adjusting the periodic time interval based on the determined changes in the sample.

However, DeVries et al. do teach

adjusting the periodic time interval based on the determined changes in the sampled information (col. 6 lines 10-24).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong in view of Digalakis et al. speaker identification because DeVries et al. teach that would produce noise tracking system that determines the effective time window in real time, so as to adapt to environmental changes in noise. (DeVries, col. 6 lines 18-23).

As to claim 9 is directed toward a computer program with a computer readable program code to implement or execute the method of claim 1, and is similar in scope and content of claim 1, therefore, claim 9 is rejected under similar rationale.

As to claim 10, which depends on claim 9, Gong teaches

instructions for periodically determining a new transducer model (col. 5 lines 24-34).

As to claim 11, which depends on claim 10, Gong teaches

the parameters of the background model are determined based on a first sample period (Fig. 2; sample period for background noise is determined before speech utterance)

the parameters of the transducer model are determined based on a second sample period (col. 5 lines 20-30; sample period for transducer model takes place during one-time adaptation (calibration), which takes place before on-line adaptation and thus inherently requires a second, distinct sampling)

As to claim 12, which depends on claim 10, Gong teaches

instructions for saving at least one of the background model (background noise is recorded and estimated, col. 2 lines 43-44 and col. 5 lines 24-34).

Claim 13 directed toward a computer readable storage medium with a computer readable program code to implement or execute the method of claim 1, and is similar in scope and content of claim 1, therefore, claim 13 is rejected under similar rationale.

Claim 14 is directed toward a computer readable storage medium with a computer readable program code to implement or execute the method of claim 1, and is similar in scope and content of claim 1, therefore, claim 14 is rejected under similar rationale.

As to claim 15, which depends on claims 1, Gong teaches

repeatedly determining the parameters of the transducer model (col. 5 lines 28-34).

As to claim 16, which depends on claim 5, Gong teaches

the transducer model estimation circuit (necessary circuit in recognizer, col. 5 lines 24-32 and col. 1 line 15 and 31-34) is configured to repeatedly determine the transducer model at the periodic time interval (Fig. 2 0.3 delay, col. 2 lines 35-45).

As to claim 17, which depends on claim 13, Gong teaches

repeatedly determining the parameters of the transducer model (col. 5 lines 25-34).

As to claim 18, which depends on claim 14, Gong teaches

determining the parameters of the transducer model (col. 5 lines 28-34).

Gong in view of Digalakis et al. does not explicitly teach adjusting the periodic time interval based, at least in part, on the collected first sampled information.

However, DeVries et al. do teach

adjusting the periodic time interval at least in part, on the collected first sampled information (col. 6 lines 10-24; DeVries et al. would necessarily use the first sampled information in a real-time application in order to readily determine the noise level changes which are analyzed using the forgetting factor in order to readily adapt to the changes in noise level).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong 's speech in view of Digalakis et al. speaker identification because an artisan of ordinary skill in the art would produce a noise tracking system that determines the effective time window in real time, so as to optimally predict the noise power for the next frame because in an automobile environment, passing cars or the shifting of gears may introduce short-term non-stationary noise. (DeVries, col. 6 lines 18-23).

As to claim 19, which depends on claim 19, Gong teaches

interval of sample (Fig. 2).

Gong in view of Digalakis et al. does not explicitly teach adjusting the length of the intervals.

However, DeVries et al. do teach

adjusting the length of the first periodic intervals based, at least in part, on a frequency (amplitude-frequency product, energy, room noise and speech, noise update speech frame, forgetting factor predict noise power) of determined changes successively sampled ones of the first sampled information (adapt real time, forgetting factor, to predict noise power for the next frame, col. 8 lines 2-6, 21-24; col. 5 lines 48-51, col. 6 lines 2-4, 10-11, 20-23).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong in view of Digalakis et al. speaker identification because an artisan of ordinary skill in the art would adjust interval of frequency samples, so as to optimally predict the noise power for the next frame. (DeVries, col. 6 lines 18-23).

As to claim 22, which depends on claim 1, Gong teaches

wherein the user identifier is based on rules associated with a phone of the speaker and a time (col. 2 lines 11-26 and Fig. 2).

As to claim 24, which depends on claim 5, Gong teaches

wherein the user identifier is based on rules associated with a phone of the speaker and a time (col. 2 lines 11-26 and Fig. 2).

5. Claims 2, 4, 6, 8, and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gong (6,418,411), in view of Digalakis et al. (5,864,810) in further view of DeVries (6,289,309), as applied to claim 1, and in further view of Thrasher et al. (2002/0052742).

As to claims 2, which depends on claim 1, Gong teaches

speech recognition modeling (Fig. 1 element 21).

Gong in view of Digalakis et al. in further view of DeVries do not explicitly teach confidence score to generate word lattices.

However, Thrasher et al. do teach

generating a confidence score (confidence measure, col. 3, paragraphs 0035-0036, Fig. 2, element 110) to determine whether the generated word lattices (page 3 paragraph 36) are acceptable (identifiers indicating which patterns may have been improperly identified, col. 3, paragraphs 0035-0036; acoustical score that measures the "acceptability" of word lattices).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify Gong in view of Digalakis et al. in further view of DeVries et al.'s noise speech enhancement such that it generates a confidence score, because an artisan of ordinary skill in the art would identify proper patterns that would provide an accurate recognizer. (Thrasher et al., col. 3, paragraph 0035).

As to claim 4, which depends on claim 2, Gong teaches

saving at least one of the parameters of the background model and the transducer model (background noise is recorded and estimated, col. 2, lines 43-44; and col. 5 lines 24-34).

As to claim 6, which depends on claim 5, Gong teaches

speech recognition modeling (Fig. 1 element 21).

Gong in view of Digalakis et al. in further view of DeVries do not teach confidence score to determine lattices.

However, Thrasher et al. do teach

generating a confidence score (confidence measure, col. 3, paragraphs 0035-0036) after applying speech recognition model (language model, Fig. 2, element 110) to determine whether the lattices (page 3 paragraph 36) are acceptable (identifiers indicating which patterns may have been improperly identified, col. 3, paragraphs 0035-0036; acoustical score that measures the "acceptability" of word lattices).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong in view of Digalakis et al. in further view of DeVries et al.'s noise speech enhancement because an artisan of ordinary skill in the art would generate a confidence score to avoid poor recognition quality. (Thrasher et al., col. 3, paragraph 0035).

As to claim 8, which depends on claim 6, Gong teaches

saving at least one of the parameters of the background model and the transducer model (background noise is recorded and estimated, col. 2, lines 43-44; and col. 5 lines 24-34).

determining the adaptation speech recognition model (adaptation of HMM for speaker and acoustic environment, col. 1, lines 38-40) based on at least one of the background model (background model is determined based on the samples taken during the sample period, col. 2 lines 43-45 & element 18, Fig. 1).

As to claim 20, which depends on claim 14, Gong teaches speech recognition (Fig. 1).

Gong in view of Digalakis et al. do not explicitly teach confidence scoring.

However, Thrasher et al. do teach

generating a confidence score after applying the speech recognition model to determine whether the generated word lattices are acceptable (confidence measure based on probable sequences provided as a result of lattice, lattice have a lexical word, in recognized speech and acoustic score, page 3 paragraphs 34-36);

comparing the confidence score to a predetermined value (page 3 paragraphs 32 and 35-36 and page 4 paragraph 40; user predetermines the value of the confidence score via listening to the results, user does comparison); and

repeating automatic speech recognition (re-launch) of the received voice request based, at least in part, on a result of the comparing of the confidence score with the predetermined value (edit recognition of speech, user re-launches application, reinitializes hypothesis, page 4 paragraph 40; user edits to reinitialize hypothesis if there is a problem with confidence score and the predetermined value).

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify Gong in view of Digalakis et al. in further view of DeVries et al.'s speech recognition model to produce Thrasher et al.'s N-best alternatives in speech recognition because an artisan of ordinary skill in the art would produce an engine that is never considering more than a predetermined maximum number of sub-paths, allowing for quicker processing (Thrasher et al., page 1 paragraph 9).

6. Claims 3 and 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gong (6,418,411), in view of Digalakis et al. (5,864,810) and DeVries (6,289,309), in view of Thrasher et al. (20020052742), as applied to claims 2 and 6, and in further view of Waibel et al. (5,712,957).

As to claim 3 which depend on claim 2, Gong teaches

the parameters of the background model are determined based on a first sample period (sample period for background noise is determined before speech utterance, Fig. 2);

the parameters of the transducer model are determined based on a second sample period (sample period for transducer model takes place during one-time adaptation (calibration), which takes place before on-line adaptation and thus inherently requires a second, distinct sampling, col. 5, lines 23-28)

Gong in view of Digalakis et al. and in further view of DeVries do not teach comparing confidence scores to determine weather to perform the ASR process again.

However, Waibel et al. do teach

the confidence score is compared to a predetermined value (threshold value) in order to determine weather to perform the automatic speech recognition process again (repeat again, col. 1, lines 56-59).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to modify Gong in combination with the speech recognition systems of Digalakis et al. and DeVries into Thrasher's method so that the confidence score is compared to a predetermined threshold value to repair misrecognition of speech. (Waibel col. 1, lines 9-12).

Claim 7 is directed toward a system with a controller to implement or execute the method of claim 3, and is similar in scope and content of claim 3, therefore, claim 7 is rejected under similar rationale.

7. Claims 21, 23 and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gong (6,418,411), in view of Digalakis et al. (5,864,810) in further view of DeVries (6,289,309), as applied to claims 1, 5 and 14, and in further view of Comerford et al. (6,107,935).

As to claim 21, which depends on claim 1, Gong teaches

user identification (col. 3 lines 5-20)

Gong in view of Digalakis et al. in further view of DeVries do not explicitly teach the identifier comprises a calling phone number.

However, Comerford et al. do teach

wherein the user identifier comprises a calling phone number (col. 11 lines 64-67 and col. 12 lines 1-20).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to implement Comerford et al.'s calling phone identifier into the method of Gong in view of Digalakis et al. in further view of DeVries because an artisan of ordinary skill in the art would have allowed for successfully verified calls; when the requesting speaker is not verified, the name and number is flagged and saved, but not placed (Comerford et al. col. 11 lines 64-67 and col. 12 lines 1-20).

Claim 23 is directed toward a system with a controller to implement or execute the method of claim 3, and is similar in scope and content of claim 3, therefore, claim 23 is rejected under similar rationale.

Claim 25 is directed toward a method to implement or execute the method of claim 3, and is similar in scope and content of claim 3, therefore, claim 25 is rejected under similar rationale.

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. See attached PTO-892.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Myriam Pierre whose telephone number is 571-272-7611. The examiner can normally be reached on 8:30-5:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MP 12/13/2005

Vijay B Chawan

VIJAY CHAWAN
PRIMARY EXAMINER